American Journal of Epidemiology
Published by Oxford University Press on behalf of the Johns Hopkins Bloomberg School of Public Health 2022.

Vol. 191, No. 12
https://doi.org/10.1093/aje/kwac154
Advance Access publication:
August 25, 2022

Practice of Epidemiology

Opioid Use Disorder Among Ohio's Medicaid Population: Prevalence Estimates From 19 Counties Using a Multiplier Method

Nathan J. Doogan*, Aimee Mack, Jianing Wang, Dushka Crane, Rebecca Jackson, Mary Applegate, Jennifer Villani, Redonna Chandler, and Joshua A. Barocas

* Correspondence to Dr. Nathan J. Doogan, Ohio Colleges of Medicine Government Resource Center, The Ohio State University, 150 Pressey Hall, 1070 Carmack Road, Columbus, OH 43210 (e-mail: Nathan.Doogan@osumc.edu).

Initially submitted June 22, 2021; accepted for publication August 19, 2022.

The decades-long overdose epidemic in the United States is driven by opioid misuse. Overdoses commonly, although not exclusively, occur in individuals with opioid use disorder (OUD). To allocate adequate resources and develop appropriately scaled public health responses, accurate estimation of the prevalence of OUD is needed. Indirect methods (e.g., a multiplier method) of estimating prevalence of problematic substance-use behavior circumvent some limitations of household surveys and use of administrative data. We used a multiplier method to estimate OUD prevalence among the adult Medicaid population (ages 18–64 years) in 19 Ohio counties that are highly affected by overdose. We used Medicaid claims data and the US National Vital Statistics System overdose death data, which were linked at the person level. A statistical model leveraged opioid-related death rate information from a group with known OUD to estimate prevalence among a group with unknown OUD status given recorded opioid-related deaths in that group. We estimated that 13.6% of the total study population had OUD in 2019. Men (16.7%) had a higher prevalence of OUD than women (11.4%), and persons aged 35–54 had the highest prevalence (16.7%). Our approach to prevalence estimation has important implications for OUD surveillance and treatment in the United States.

indirect prevalence estimation; opioid use disorder; prevalence

Abbreviations: ACS, American Community Survey; CrI, credible interval; HCS, Helping to End Addiction Long-term (HEALing) Communities Study; ICD-10, *International Classification of Diseases*, *Tenth Revision*; MOUD, medication for opioid use disorder; OUD, opioid use disorder.

BACKGROUND

The United States continues to experience a decades-long exponential growth of per-capita overdose deaths (1), constituting an epidemic. Opioids were associated with nearly 70% of fatal overdoses in 2018 (1, 2), making them the most prominent substance underlying the epidemic today. Some of the most vulnerable people—including people who are unemployed or experiencing homelessness, those with lower educational attainment, those who lack social support and access to health insurance, and people with involvement in the legal system—are overrepresented among those who die from opioid overdoses (3, 4). Thus, along with availability of opioids and lack of access to medications for opioid use disorder (OUD), relative vulnerability is likely to be a key driver of the epidemic. Resources have now been dedicated to solving the opioid crisis using evidence-based interventions (e.g., Chandler et al. (5) and Kerr (6)). Accurate monitoring of the incidence and prevalence of OUD with sufficiently high geographic resolution and across important demographic levels is challenging yet critical for resource allocation and determination of intervention effectiveness. Self-report survey methods, which are often considered the gold standard for measuring health behaviors, are known to underestimate problematic drug use behavior (7). As such, alternative indirect approaches to estimating the size of hidden populations have become popular (8–11). The purpose of this study was to estimate OUD prevalence among Medicaid participants—who, by virtue of their relative vulnerability, are likely to be at elevated risk of

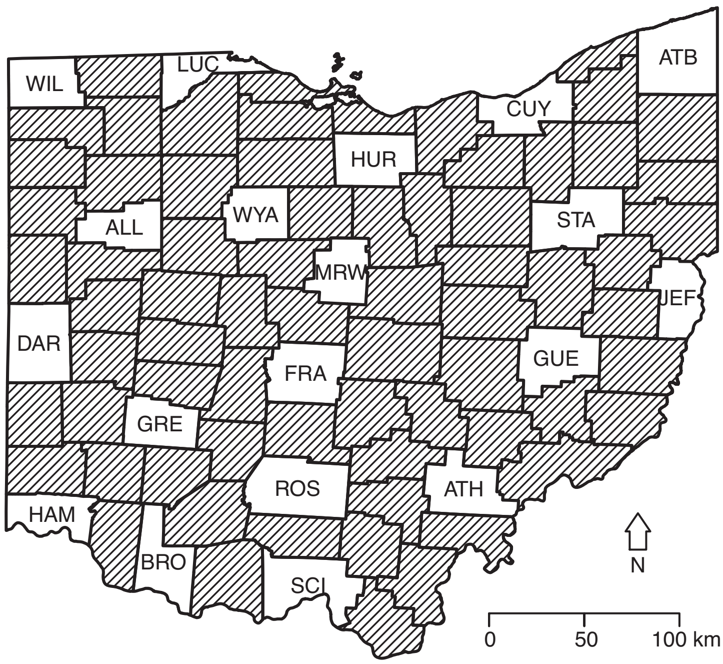

Figure 1. The 19 Ohio counties in the Helping to End Addiction Long-term (HEALing) Communities Study (initiated in 2020), indicated in solid white, labeled with an abbreviation of the county name: ALL, Allen; ATB, Ashtabula; ATH, Athens; BRO, Brown; CUY, Cuyahoga; DAR, Darke; FRA, Franklin; GRE, Greene; GUE, Guernsey; HAM, Hamilton; HUR, Huron; JEF, Jefferson; LUC, Lucas; MRW, Morrow; ROS, Ross; SCI, Scioto; STA, Stark; WIL, Williams; WYA, Wyandot.

OUD—within the state of Ohio in strata defined by county, age, and sex, using a flexible indirect method known as a multiplier method.

Crisis and response

Opioids were associated with 46,802 (69.5%) of the total 67,367 fatal overdoses recorded in the United States in 2018 (2). The Society of Actuaries has estimated a total economic burden for the United States, inclusive of health care, mortality, criminal justice, child and family assistance, education, and lost productivity costs, to be $179.4 billion in 2018 alone (12). While no current estimate was found, a 2013 estimate of the cost just to US Medicaid systems was approximately $8 billion (in 2017 dollars), inclusive only of services provided for OUD and excess costs in other services for people with OUD (13). Citing the estimated 11.5 million Americans misusing prescription opioids, a continual rise in overdose deaths, and one-third of all child removals being attributed to parental substance use, the United States declared a public health emergency in 2017 (14).

We focus on one federal response to the current state of the overdose crisis that highlighted the need to estimate prevalence of OUD in Ohio more accurately. The Helping to End Addiction Long-term (HEALing) Communities Study (HCS), an ongoing study initiated in 2020, is a multisite, parallel-group, cluster-randomized, wait-list controlled trial focused on 67 US communities in 4 US states—Kentucky, Massachusetts, New York, and Ohio (5). Nineteen of those communities were Ohio counties (see Figure 1)—which were carefully selected to include communities with high overdose death rates and to optimize intervention effectiveness and evaluation quality. The study objectives were to use data-driven approaches to inform selection and implementation of evidence-based practices where they are most needed to reduce opioid overdose deaths (goal: 40% reduction) and rigorously evaluate the intervention. Accurate estimation of OUD prevalence at a county level and for demographic subsets of the population could benefit both objectives.

Challenges to direct estimation of OUD prevalence

Estimating the prevalence of OUD and other problematic drug use behavior with household survey methods is challenging for multiple reasons (7). The stigma associated with problematic drug use (15) can inhibit self-reports, leading to a social desirability bias. Additionally, the behavior is relatively rare, and individuals who experience it are more likely to be unstably housed, incarcerated, or hospitalized, which further challenges precise estimation, particularly at substate geographic levels, for national surveys. Therefore, household survey-based estimates are expected to be considerably downward-biased (16) and available only for large geographic areas.

Estimating prevalence of OUD from administrative data, such as Medicaid claims data, is also a challenge. People with OUD may not be identified by the health-care system.

This could result from a lack of health service utilization necessary for OUD to be identified, incomplete data regarding services or diagnoses, a preference for privacy by the patient, or an unwillingness of the provider to apply a diagnosis without commitment to treatment from the patient.

An indirect method to estimate OUD prevalence

A number of alternative approaches to prevalence estimation—classified as indirect methods—have become popular (7). One of these is referred to as a multiplier method.

For our purposes, it is based on a model in which the number of opioid-related deaths d during a year is the product of the yearly rate of opioid-related death λ for people with OUD and the total number of person-years at risk of opioid-related death (i.e., with OUD) η. Person-years at risk can be thought of as the number of people with OUD at any time during the year-long study.

$$\lambda\eta = d$$

If d is known (e.g., via Vital Statistics data) and if λ can be estimated for a representative group, the remaining unknown value η can be inferred. Given an estimated rate of opioid-related death ($\hat{\lambda}$), we want an estimate ($\hat{\eta}$) of the number of people who must have had OUD to result in the observed number of these deaths (d). An algebraic rearrangement reveals a solution—the quotient of the number of opioid-related deaths (dividend) and the estimate of the opioid-related death rate (divisor):

$$\hat{\eta} = \frac{d}{\hat{\lambda}}$$

This basic approach constitutes the essence of the procedures used in the present study.

Study purpose

The purpose of the study was to estimate the prevalence of OUD among Medicaid enrollees aged 18–64 years in 19 Ohio HCS counties. We adapted the approach of Jones et al. (17), who constructed a statistical model that amounts to a flexible multiplier method for identifying the size of a group of people with latent OUD given data about known diagnoses, treatments, and opioid-related deaths. We report on the model-building process and the estimates of prevalence. We then discuss the results in the context of Ohio's overdose epidemic and some considerations for further use of this approach that may be relevant to the accuracy of our estimates and to other researchers or policy makers who may wish to replicate this work.

METHODS

Data sources and linkage

The data used for this study were derived from 2 primary sources: 1) Medicaid administrative data relevant to OUD diagnosis and medication treatment, and 2) Ohio Vital Statistics data about opioid-related overdose deaths. Two additional sources that supported estimation, but were not critical to the approach, were 3) the American Community Survey (ACS), and 4) a continuous measure of rurality known as geographic isolation (18). The administrative data represented the period January 1 through December 31, 2019. The ACS data were 5-year averages ending in 2019. The geographic isolation measure was calculated based on 2017 US Census estimates and auxiliary data from no earlier than 2010. Inclusion criteria were Medicaid enrollment at any point during the study period, age between 18 and 64 years, and residence in one of the 19 Ohio HCS counties according to the most recent Medicaid enrollment record for the individual.

Medicaid administrative data were used to identify county of residence, sex (male or female), age group (years: 18–34, 35–54, 55–64), OUD diagnosis (observed OUD, unknown OUD status), and monthly treatment status. A diagnosis of OUD was measured as at least 1 encounter resulting in an *International Classification of Diseases*, *Tenth Revision* (ICD-10), diagnosis code of F111 (opioid abuse) or F112 (opioid dependence) in any setting, appearing in any diagnosis position on the claim during the study period. An individual was considered "treated" during a month if they were associated with at least 1 claim for medication for OUD (MOUD) treatment—inclusive of methadone, buprenorphine, or naltrexone—during the month. They were "untreated" otherwise, including if they received behavioral treatment or if their MOUD was paid by an entity other than Medicaid, constituting a limitation of the available data. Months unenrolled from Medicaid were counted as untreated months unless the person died for any reason, in which case the months following death were not counted as treated or untreated.

Vital Statistics were used to identify opioid-related fatal overdoses during the study period. The death certificates must have reported an underlying cause of drug poisoning with ICD-10 codes in the ranges X40–X44 (unintentional) or Y10–Y14 (undetermined intent), as well as a cause of death code of T40.0 (opium), T40.1 (heroin), T40.2 (natural opioid analgesics), T40.3 (methadone), T40.4 (synthetic opioid analgesics other than methadone), or T40.6 (other and unspecified narcotics). This definition deviated from that of Jones et al. (17) in that we did not focus on ICD-10 F11 and F19 codes, which are used inconsistently in the United States.

To check the sensitivity of our estimates to our deviation from Jones et al. (17), we performed a check in which we changed the definition of opioid-related deaths to be deaths with a contributing cause of death code F11 or F19, and where an F19 code must also appear with a T code in the range T40.0–T40.4 or T40.6.

Medicaid and Vital Statistics data sources were linked at the individual level using the Social Security Number, which was available in both and expected to be recorded with high fidelity.

Analytical data set

Once the data were linked, an analytical data set was constructed to characterize each individual stratum. Each stratum was characterized by the following: 1) stratifying characteristics county, sex, and age group; 2) the total size of the stratum population; 3) the number with observed OUD; 4) the cumulative person-years not on MOUD (number of months divided by 12); 5) the number of opioid-related deaths that occurred while off MOUD treatment among those with an observed OUD; and finally, 6), the number of opioid-related deaths among those with an unknown OUD status.

A priori, variability in opioid-related death rate and prevalence across counties is plausible. The model could capture this with county-specific fixed or random effects. However, if variations are correlated with observable county characteristics, a more parsimonious approach could be to add measurements of those characteristics to the model. We added 3 county-level variables to the analytical data set for consideration in model selection. Two of them were county summaries from the ACS: the proportion of the full county population that 1) was people of color and 2) had a household income below 150% of the federal poverty level. The third was a continuous measure of tract-level rurality, called geographic isolation, aggregated to the county level using a population-weighted mean (18).

Strategy

We used the analytical data set to estimate both the opioid-related death rate and the total population with OUD. We express this as a 2-stage process for clarity. In practice, they were accomplished simultaneously when the model was fitted to the data. A more technical description of the model can be found in Web Appendix 1 (available at https://doi.org/10.1093/aje/kwac154).

First, we used the number of opioid-related deaths while untreated among those with OUD and the number of months spent untreated to estimate the rate of opioid-related death while untreated among those with OUD. Second, among those with no OUD diagnosis (i.e., OUD status is unknown), we estimated the number with latent OUD by leveraging the estimated death rate (while untreated), the number of opioid-related deaths, and the model of deaths at the core of our approach.

The use of the estimated opioid-related death rate implied that we assumed people with a latent OUD were untreated and therefore that they died at the same rate as people with observed OUD while untreated. The sum of the estimated number with a latent OUD and the number with observed OUD constituted our estimate of the total with OUD.

Noting that the prevalence estimate depended critically on the estimated opioid-related death rate, and the death rates were likely to vary due to characteristics of the individual or where they lived, we expected that our approach would benefit from flexibility in this regard. To achieve this, we disaggregated the data into a set of strata defined by county, age group, and sex. This facilitated stratum-specific death rate estimates, which were expected to enhance the accuracy of stratum-specific and aggregated prevalence estimates.

Independently applying this approach to individual strata is likely to be impractical given that some strata have very few or zero opioid-related deaths during the study time period, thus leading to uncertain or unidentifiable estimates. We adapted an approach used in the United Kingdom by Jones et al. (17) that selectively pools information from multiple strata while still estimating stratum-specific prevalence. The level and type of pooling was determined by a model selection approach that balances the bias-variance trade off.

Model selection and reporting

The analysis consisted of an iterative model selection approach. The overarching pattern was to fit sensible models by starting simple and looking for ways to improve upon them (in terms of a complexity-penalized criterion) by increasing the complexity using theoretical and empirical insights by looking at model fit via posterior predictive checks (19). We restricted the space of potential models to those with the same specification between the death and prevalence components of the model because model fitting is time-intensive. Our approach included models that used regularizing priors on maximally complex models with stratum-specific parameters (e.g., least absolute shrinkage and selection operator (LASSO) (20) and the horseshoe prior (21)). While we believe finding a well-fitted model to be imperative, we do not wish to advocate for any particular model selection approach.

Relative fit was measured using the leave-one-out information criterion (LOOIC), which balances model fit to the data with parsimony thus controlling over-fit to the data (22). To examine absolute fit, we compared the distribution of opioid-related death counts in the data with those generated by the fitted model. We used a χ^2 test of the null hypothesis that the observed count frequencies arise from the same joint distribution that the model implies. The test statistic was a Mahalanobis distance due to correlation among the death count frequencies. For example, variations among predictions in the frequency of strata with 5 deaths is likely to be correlated with the frequency of strata with death counts of 4 and 6. We were satisfied with absolute fit when the null hypothesis was not rejected based on a threshold of $P = 0.05$. Unreported, but helpful with model selection, was visual examination of observed versus predicted opioid-related death counts in each individual stratum to learn in which cases a model was misfitting, which then combined with domain knowledge to guide next steps. The selected model had acceptable absolute fit and a favorable LOOIC value relative to alternatives.

We report the selected model specification (see Web Appendix 2 and Web Table 1 for alternative specifications and a comparison of fit statistics), results of the fit assessment, and aggregate and various disaggregated estimates of prevalence. We also report the prevalence estimate generated by a model fit to check the sensitivity of our estimates to our definition of opioid-related deaths, which differed from the one used by Jones et al. (17). All reported estimates are provided with either visual or numeric representations of a 95% credible interval (CrI).

The data analysis was completed within the R statistical computing environment (R Foundation for Statistical Computing, Vienna, Austria (23)), and the models were fitted using the R2Jags package, an R interface to the JAGS

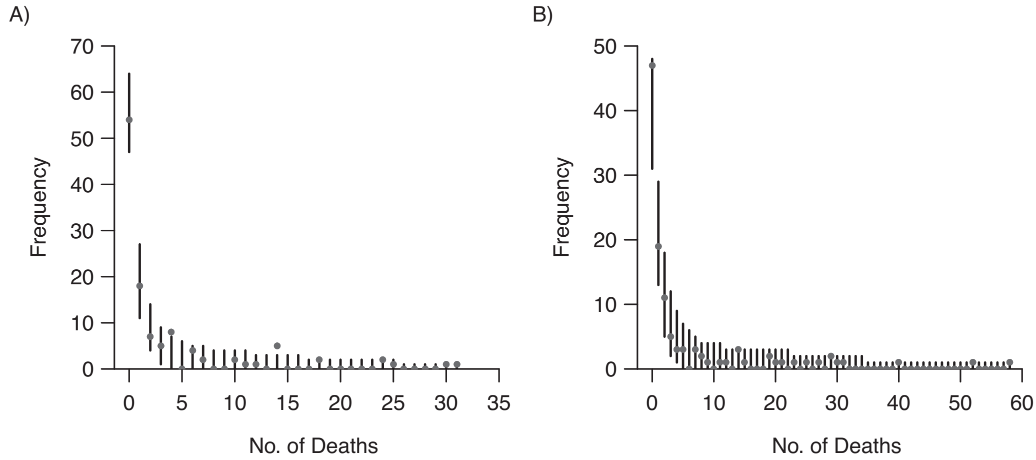

Figure 2. Comparison of the observed frequency of opioid-related death counts with the 95% posterior prediction interval of the predicted frequency of death counts generated by the fitted model, using data from Ohio, 2019. A) χ^2 test resulting in $P = 0.23$; B) and χ^2 test resulting in $P = 0.70$. These values indicate that the data could have been generated by a process similar to what our model represents.

probabilistic programming language (24, 25). We simulated 250,000 samples from the posterior distribution and reviewed trace plots to ensure the process was stationary and resulted in sufficient exploration of the posterior distribution.

RESULTS

Selected model specification

The optimal model by our selection process included age and sex as fixed effects and county-level random effects in both main model components. No attempted specification that augmented the selected model with fixed effects or replaced random effects with 1 or more county-level fixed effects (i.e., no combination of poverty, people of color, or rurality) was better according to both of our criteria (favorable LOOIC and suitable absolute fit).

Model fit

Figure 2A and 2B visualize model goodness-of-fit to the data. The test of the null hypothesis, that the observed counts came from a process like that described by the model, is not rejected for either set of death counts (both $P > 0.05$). Thus, the selected model had acceptable absolute fit to the data.

OUD prevalence estimates

We estimated that 116,584 (13.6%, 95% CrI: 12.7, 14.7) of the study population had an OUD during the 2019 calendar year. This is compared with a naive estimate based on direct observation of OUD diagnoses in Medicaid administrative data, which is 53,508 (6.3%; see Table 1).

By our model, men (16.7%, 95% CrI: 15.3, 18.3) were more likely to have an OUD than women (11.4%, 95% CrI: 10.3, 12.6). The age group 35–54 years had a higher point estimate (16.7%, 95% CrI: 15.3, 18.4) than the age group 18–34 years (10.8%, 95% CrI: 9.7, 12.0) and a somewhat higher estimate compared with the age group 55–64 years (14.5%, 95% CrI: 11.8, 18.0). Visual comparisons are available in Figure 3.

Table 1 contains county-specific and the overall estimates of OUD prevalence along with 95% CrIs and other useful information about counties, such as the total Medicaid population size (ages 18–64 years in study counties), the death counts for both groups, and the naive OUD prevalence estimates that were direct counts of Medicaid enrollees with observed OUD diagnoses.

Table 2 contains stratum-specific estimates of the percentage of the study population with OUD as estimated by our model. It presents prevalence estimates for all 114 strata.

Sensitivity analysis

We checked the sensitivity of our methods to our definition of opioid-related deaths. When we used a definition similar to that of Jones et al. (17), we identified 487 deaths total, making it a much more restrictive definition (compared with 1,045 deaths from our definition). The model was the same one we selected for our data, and it had an acceptable absolute fit with the data for the sensitivity check (P for death rate model fit test = 0.798; P for prevalence model fit test = 0.492). Despite our expectation that the more restrictive definition would lead to a more conservative prevalence estimate, this model resulted in an overall point estimate of 15.7% (95% CrI: 12.7, 22.7)—2.1 percentage points higher than the overall prevalence we reported.

DISCUSSION

Major findings

As the US overdose epidemic continues and as resources are consumed to combat it, accurate surveillance

Table 1. County-Level Descriptive and Inferential Statistics on Prevalence Estimates for Opioid Use Disorder Among the Medicaid Population (Aged 18–64 Years), Using Claims Data, Ohio, 2019

Country	Total Medicaid Population		Naive Total With OUD		Opioid-Related Deaths		Estimated Total With OUD		Estimated Percentage With OUD	
	No.	% of Country Population	No.	% of Medicaid Population	With Observed OUD	With Unobserved OUD Status	No.	95% CrI	%	95% CrI
Franklin	201,222	24.0	15,102	7.5	113	171	30,746	27,398, 34,761	15.3	13.6, 17.3
Cuyahoga	238,511	31.2	8,231	3.5	86	164	20,200	17,284, 23,735	8.5	7.2, 10.0
Hamilton	124,412	24.7	7,279	5.9	63	116	15,919	13,554, 18,924	12.8	10.9, 15.2
Lucas	82,150	31.2	5,543	6.7	53	56	10,104	8,689, 11,948	12.3	10.6, 14.5
Scioto	18,749	40.8	3,149	16.8	21	38	6,993	5,596, 8,995	37.3	29.9, 48.0
Stark	55,714	25.2	2,946	5.3	20	24	6,447	4,912, 8,809	11.6	8.8, 15.8
Ross	16,629	34.6	1,996	12.0	8	9	3,624	2,979, 4,856	21.8	17.9, 29.2
Greene	17,738	17.0	1,319	7.4	10	15	3,085	2,166, 4,513	17.4	12.2, 25.4
Jefferson	13,103	33.1	1,327	10.1	6	5	2,915	2,195, 4,225	22.2	16.8, 32.2
Ashtabula	19,032	32.8	1,147	6.0	9	8	2,734	2012, 4,087	14.4	10.6, 21.5
Allen	16,609	27.0	861	5.2	1	5	2,711	1,499, 4,859	16.3	9.0, 29.3
Athens	10,098	21.1	1,215	12.0	1	0	2,257	1779, 3,240	22.4	17.6, 32.1
Huron	8,457	24.6	647	7.7	3	8	2,047	1,281, 3,338	24.2	15.1, 39.5
Brown	7,539	29.4	967	12.8	2	5	2,001	1,361, 3,175	26.6	18.1, 42.1
Guernsey	7,386	32.1	821	11.1	1	3	1,792	1,282, 2,845	24.3	17.4, 38.5
Morrow	4,667	22.1	364	7.8	2	2	965	600, 1,641	20.7	12.9, 35.2
Williams	4,589	21.2	183	4.0	0	3	882	427, 1,709	19.2	9.3, 37.2
Darke	5,797	19.8	266	4.6	1	0	741	420, 1,421	12.8	7.2, 24.5
Wyandot	2,174	17.0	145	6.7	0	0	410	251, 747	18.9	11.5, 34.4
Total	854,576	27.0	53,508	6.3	400	632	116,584	108,929, 125,248	13.6	12.7, 14.7

Abbreviations: OUD, opioid use disorder; CrI, credible interval.

that facilitates targeted resource allocation and outcome-monitoring is increasingly important. We used an indirect approach—a multiplier method—to estimate the prevalence of OUD among persons aged 18–64 years, in Ohio's Medicaid population, among 19 counties experiencing high impact from it. Prior to our study, the most widely used estimate of OUD prevalence in Ohio was 1.0% for Ohioans aged 12 or older and 1.9% among only the young adult population (18–25 years) (26). Using the multiplier method, we estimated an approximately 13.6% prevalence among our study population. Notably, this estimate is high relative to other uses of related indirect methods (27), and a substantial portion of the difference can be explained by our focus on the Medicaid population. Nevertheless, the result is striking. The development of our study included a review of the findings by local experts of the opioid crisis. Equally striking was their lack of surprise with the overall estimate, which lends a modicum of face validity to our estimates. And in conjunction with widely used prevalence estimates between 1% and 2%, it also suggests that the general public is likely to have a limited understanding of the extent of the overdose epidemic and who bears the greatest impact of it. Estimates that minimize the problem are unlikely to result in an appropriate and equitable response.

The sex differences in prevalence that we found were similar to those identified in national surveys, with men more likely to have OUD than women. Nevertheless, the overall prevalence for women—including those of reproductive age, for whom the deleterious potential could affect more than one life—was high. In addition to facilitating targeted general OUD interventions, our stratum-specific estimates could be used to target interventions aimed specifically at reducing the incidence of neonatal abstinence syndrome.

Our analysis also suggests that in Ohio, the middle age group (35–54 years) has the highest prevalence of OUD. This differs from the relative prevalence patterns observed in the general US population, where the young adult group appears to have the highest prevalence (28). This may signal unique characteristics of the Ohio Medicaid population worthy of further study.

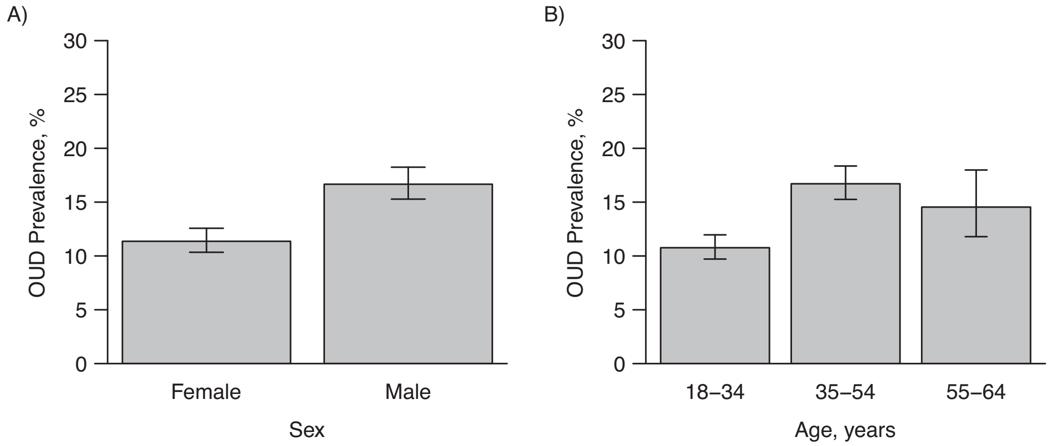

Figure 3. These plots display opioid use disorder (OUD) prevalence and 95% credible intervals among the Medicaid population aged 18–64 years in the 19 Ohio counties in the Helping to End Addiction Long-term (HEALing) Communities Study, 2019. The overall estimate is 13.6% (95% credible interval: 12.7, 14.7). A) Prevalence by sex; B) prevalence by age group.

Generalization of our overall estimates for the Medicaid population (ages 18–64 years) in 19 HCS counties to all counties in the state is complicated by the nature of the HCS county selection process. The study focused on counties meeting a minimum rate of overdose deaths, which suggests the selected counties may have higher prevalence than nonselected counties.

Considerations for further use of multiplier methods for this purpose

Key assumptions. This method of estimating OUD prevalence critically relies on a valid estimate of the opioid-related death rate used to estimate OUD prevalence in the group with unknown OUD status. We estimated the opioid-related death rate using data from a group with a known OUD status. By definition in our study, the unknown-OUD group is untreated, and so the estimated opioid-related death rate had to account for treatment (i.e., not represent a death rate during treatment). More generally, anything that might create a difference in the death rates between the 2 groups can bias the prevalence estimate. For example, it could be that people with an OUD diagnosis are more likely to have positive influences in their lives that led them to diagnosis. Such latent support may decrease the likelihood of death, which could imply that we underestimated the death rate we applied to the group without a diagnosis, and therefore overestimated the prevalence.

A second assumption related to the off-treatment opioid-related death rate was that we defined "treatment" to be exclusive of nonmedication behavioral health (BH) treatments as well as MOUD treatments that were paid for by a non-Medicaid entity. What we called "untreated" might have included genuinely beneficial treatments. If so, our estimates of the "untreated" death rate may have been too low, causing an upward bias of prevalence.

Definition of opioid-related death. We chose to include opioid-related overdose deaths using a standard and inclusive set of diagnostic codes. In the United States, the data suggest that the majority of opioid-related overdoses occur among people with OUD (mild, moderate, or severe, as defined by the *Diagnostic and Statistical Manual of Mental Disorders*, *Fifth Edition*) (29–31). We recognize that this standard may be different in other countries and depends heavily on the epidemiology of substance use. In England, for example, ICD-10 codes have been restricted to those using only F11 and F19 roots (17).

We checked the sensitivity of our prevalence estimates to our definition of opioid-related deaths by trying a definition closer to that of Jones et al. (17). We found the prevalence estimates increased somewhat with the more restrictive definition, and we concluded that our more inclusive definition did not bias our prevalence estimates upward, as anticipated. Nevertheless, the ideal definition of opioid-related deaths for this purpose should continue to be considered carefully, and may depend on local reporting standards.

Definition of OUD in claims data. Our simple definition of OUD in Medicaid claims may have inadvertently missed persons with OUD or included persons who may not have true OUD. Other algorithms for identifying OUD in claims data exist and could be considered (e.g., Wakeman et al. (32)).

Linkage errors. Compared with the work of Jones et al. (17), our model does not include a correction for mismatches between the death and treatment/diagnosis data. In their approach, the data for the group without OUD diagnosis was not explicitly linked to death data. Instead, all deaths not linked to treatment/diagnosis data were assumed to have occurred among the group for whom OUD status was unobserved. Thus, missed linkages would erroneously reduce deaths in the known-OUD group and increase them

Table 2. Opioid Use Disorder Prevalence Among the Medicaid Population Aged 18–64 Years in the 19 Counties in the Helping to End Addiction Long-term (HEALing) Communities Study, Ohio, 2019

County and Sex	Age in Years					
	18–34		35–54		55–64	
	%	95% CrI	%	95% CrI	%	95% CrI
Allen						
Female	11.1	5.9, 20.9	16.9	9.1, 30.6	14.8	7.4, 27.6
Male	16.1	8.6, 29.5	23.7	13.2, 40.9	21.1	11.2, 37.6
Ashtabula						
Female	9.4	6.9, 14.5	14.4	10.0, 22.0	18.1	11.9, 28.1
Male	13.9	10.2, 20.9	20.7	14.8, 30.4	12.7	8.2, 20.5
Athens						
Female	15.2	12.9, 22.5	22.5	17.0, 32.8	19.8	13.6, 30.6
Male	21.5	16.2, 31.4	30.9	23.3, 43.3	27.6	19.0, 40.8
Brown						
Female	18.4	11.5, 31.5	27.0	18.1, 42.8	23.8	14.5, 40.2
Male	25.9	16.8, 42.0	35.9	24.7, 54.1	32.3	20.4, 51.5
Cuyahoga						
Female	5.4	4.3, 4.6	12.7	10.7, 15.1	11.1	8.7, 14.3
Male	8.1	6.6, 10.0	8.6	7.1, 10.4	7.4	5.6, 9.7
Darke						
Female	8.6	4.8, 17.2	13.2	6.9, 25.7	11.5	5.2, 23.7
Male	12.6	6.9, 24.3	18.8	10.1, 34.9	16.6	8.0, 32.1
Franklin						
Female	10.2	8.6, 12.1	15.7	13.4, 18.5	13.7	10.5, 17.7
Male	15.0	12.9, 17.4	22.4	19.6, 25.6	19.8	15.5, 24.9
Greene						
Female	12.0	8.1, 18.3	18.0	12.4, 26.6	15.9	9.9, 24.9
Male	17.2	11.6, 25.7	25.4	17.9, 36.1	22.5	14.5, 33.7
Guernsey						
Female	16.8	11.7, 28.4	24.7	16.6, 40.0	30.0	19.4, 47.9
Male	23.8	17.0, 38.0	33.5	23.6, 50.7	22.0	13.6, 37.3
Hamilton						
Female	8.5	6.8, 10.6	13.1	10.8, 16.1	11.5	8.7, 15.2
Male	12.5	10.3, 15.4	19.0	15.9, 22.7	16.7	12.9, 21.5
Huron						
Female	17.1	9.9, 29.8	34.0	21.9, 52.4	30.4	17.9, 49.6
Male	24.4	15.4, 40.0	25.1	15.2, 41.3	22.3	12.4, 38.6
Jefferson						
Female	15.3	11.6, 22.9	22.6	16.6, 33.2	19.9	13.1, 30.9
Male	21.6	15.6, 32.0	30.9	22.7, 43.9	27.6	18.6, 41.1
Lucas						
Female	8.1	6.6, 10.0	12.6	10.6, 15.4	10.9	8.0, 14.6
Male	12.0	10.0, 14.6	18.3	15.6, 21.6	15.9	12.0, 21.0

Table continues

Table 2. Continued

County and Sex	18–34 %	18–34 95% CrI	35–54 %	35–54 95% CrI	55–64 %	55–64 95% CrI
Morrow						
Female	14.2	8.4, 25.6	20.9	12.1, 36.2	18.4	9.7, 33.9
Male	20.2	12.5, 34.9	28.7	17.5, 46.5	25.7	14.2, 44.1
Ross						
Female	14.8	12.8, 20.3	21.9	17.2, 29.9	27.0	19.4, 37.8
Male	21.0	16.8, 28.5	30.4	24.1, 40.0	19.6	14.0, 28.1
Scioto						
Female	27.0	20.4, 36.9	37.8	30.1, 48.7	33.9	24.3, 47.3
Male	36.2	27.9, 47.5	48.3	39.6, 59.6	43.9	32.7, 58.2
Stark						
Female	7.6	5.5, 10.8	11.9	8.9, 16.5	10.4	7.1, 15.3
Male	11.3	8.3, 15.8	17.2	13.0, 23.4	15.0	10.4, 21.7
Williams						
Female	13.3	5.8, 27.7	27.4	13.7, 49.6	24.4	10.9, 46.8
Male	19.2	9.2, 37.6	19.9	9.1, 38.7	17.6	7.4, 36.1
Wyandot						
Female	13.0	7.0, 25.4	19.4	10.4, 35.8	17.1	8.2, 33.5
Male	18.9	12.3, 34.6	26.9	15.6, 46.4	23.9	12.3, 43.8

Abbreviation: CrI: credible interval.

in the unknown-OUD group. Jones et al. (17) cleverly built a correction into their model to account for this, which takes advantage of the Bayesian framework by incorporating prior beliefs about mismatch probabilities, which could not be estimated. In our case, everyone from both groups was explicitly linked with death data, and the linkage was on a high-quality linkage variable—Social Security Number. While mismatch could still bias the death rate estimate, linkage errors should be equally likely for both groups, rendering the biased estimate appropriate for use to estimate prevalence in the unknown-OUD group.

Model selection. We were fairly restrictive about our model selection approach. This was necessary because drawing Markov chain Monte Carlo (MCMC) samples for a single model was time-intensive. Despite our limited search, we were able to identify a suitable model in terms of absolute fit to the data. Therefore, we are confident that this limitation was not problematic in our case. However, others using these methods may require a less-restrictive search to find a well-fitted model.

Spatial correlation. If opioid-related death rates are spatially correlated and the available exogenous model predictors (e.g., rurality) do not explain the correlation, a hierarchical model that uses a nonindependent prior (e.g., a Gaussian process or intrinsic conditional auto-regressive (ICAR) prior) on the random effects could make more efficient use of the data by incorporating information sharing among geographically proximal areas. We did not explore this possibility because our aim was not to substantially extend the model of Jones et al. (17).

Expanding the scope. Our study was focused on the Medicaid population in Ohio. Other users may have the benefit of an all-payers database with treatment and diagnosis data covering the entire state, which would expand their scope. Alternatively, other sources of data about treatment, like a prescription drug monitoring program, could augment or replace Medicaid data and allow estimation among the general population. Even if complete information about treatment in the state is not available, estimates of death rates among the treated and untreated (separately) along with complete death data and information about the ratio of treated to untreated across the state could facilitate an estimate of prevalence. However, the latter approach could rely more dubiously on assumptions about the congruence between the group used to estimate death rates and the group with which those death rates are leveraged to estimate OUD prevalence. Nevertheless, it may be better than other available options.

CONCLUSIONS

A number of important steps need to be taken to decrease opioid-related overdoses in the United States, including expansion of treatment access, reduction of stigma associated with drug use, and prevention of drug use initiation. Simultaneously, however, we need accurate estimates of opioid use and opioid use disorder so that we can appropriately scale and target treatment and prevention interventions. Traditional epidemiologic tools for estimating prevalence of problematic drug use behavior, such as direct estimation with surveys, are insufficient. Our study used an innovative indirect multiplier method to provide OUD prevalence estimates for Ohio counties that have been particularly devastated by the opioid epidemic. We believe these estimates will directly translate into better service delivery and resource allocation. Importantly, this method can be adapted by other states and jurisdictions who are seeking to better characterize the size of their local epidemic.

ACKNOWLEDGMENTS

Author affiliations: Ohio Colleges of Medicine Government Resource Center, The Ohio State University, Columbus, Ohio, United States (Nathan J. Doogan, Aimee Mack, Dushka Crane, Rebecca Jackson); Rafik B. Hariri Institute for Computing and Computational Science & Engineering, Boston University, Boston, Massachusetts, United States (Jianing Wang); Ohio Department of Medicaid, Columbus, Ohio, United States (Mary Applegate); National Institute on Drug Abuse, National Institutes of Health, Gaithersburg, Maryland, United States (Jennifer Villani, Redonna Chandler); and Department of Medicine, University of Colorado School of Medicine, Denver, Colorado, United States (Joshua A. Barocas).

This research was supported by the National Institutes of Health through the NIH HEAL Initiative (awards UM1DA049394, UM1DA049406, UM1DA049412, UM1DA049415, UM1DA049417, and DP2DA051864). The ClinicalTrials.gov Identifier is NCT04111939.

Due to the sensitive nature of the underlying data, they are not available for distribution.

Conflict of interest: none declared.

REFERENCES

1. Jalal H, Buchanich JM, Roberts MS, et al. Changing dynamics of the drug overdose epidemic in the United States from 1979 through 2016. *Science*. 2018;361(6408):eaau1184.
2. Wilson N. Drug and opioid–involved overdose deaths—United States, 2017–2018. *MMWR Morb Mortal Wkly Rep*. 2020;69(11):290–297.
3. Altekruse SF, Cosgrove CM, Altekruse WC, et al. Socioeconomic risk factors for fatal opioid overdoses in the United States: findings from the Mortality Disparities in American Communities Study (MDAC). *PLoS ONE*. 2020; 15(1):e0227966.
4. van Draanen TC, Mitra S, et al. Socioeconomic marginalization and opioid-related overdose: a systematic review. *Drug Alcohol Depend*. 2020;214:108127.
5. Chandler RK, Villani J, Clarke T, et al. Addressing opioid overdose deaths: the vision for the HEALing communities study. *Drug Alcohol Depend*. 2020;217:108329.
6. Kerr T. Public health responses to the opioid crisis in North America. *J Epidemiol Community Health*. 2019;73(5): 377–378.
7. Hickman M, Taylor C, Chatterjee A, et al. Estimating the prevalence of problematic drug use: a review of methods and their application. *Bull Narc*. 2002;18, 15–32.
8. Sweeting MJ, De Angelis D, Ades AE, et al. Estimating the prevalence of ex-injecting drug use in the population. *Stat Methods Med Res*. 2009;18(4):381–395.
9. King R, Bird SM, Hay G, et al. Estimating current injectors in Scotland and their drug-related death rate by sex, region and age-group via Bayesian capture–recapture methods. *Stat Methods Med Res*. 2009;18(4):341–359.
10. Friedman SR, Tempalski B, Cooper H, et al. Estimating numbers of injecting drug users in metropolitan areas for structural analyses of community vulnerability and for assessing relative degrees of service provision for injecting drug users. *J Urban Health*. 2004;81(3):377–400.
11. Russell Bernard H, Hallett T, Iovita A, et al. Counting hard-to-count populations: the network scale-up method for public health. *Sex Transm Infect*. 2010;86(suppl 2):ii11–ii15.
12. Davenport S, Caverly M, Weaver A. Economic impact of non-medical opioid use in the United States. 2019;93.
13. Leslie DL, Ba DM, Agbese E, et al. The economic burden of the opioid epidemic on states: the case of Medicaid. *Am J Manag Care*. 2019;25(13):S243–S249.
14. Department of Health and Human Services. Determination that a public health emergency exists nationwide as the result of the opioid crisis (10/26/2017). https://www.phe.gov/emergency/news/healthactions/phe/Pages/opioids.aspx. Accessed September 22, 2022.
15. Corrigan P, Schomerus G, Shuman V, et al. Developing a research agenda for understanding the stigma of addictions part i: lessons from the mental health stigma literature. *Am J Addict*. 2017;26(1):59–66.
16. Murphy SM, Rosenman R. The "real" number of Washington state adolescents using marijuana, and why: a misclassification analysis. *Subst Use Misuse*. 2019;54(1):89–96.
17. Jones HE, Harris RJ, Downing BC, et al. Estimating the prevalence of problem drug use from drug-related mortality data. *Addiction*. 2020;115(12):2393–2404.
18. Doogan NJ, Roberts ME, Wewers ME, et al. Validation of a new continuous geographic isolation scale: a tool for rural health disparities research. *Soc Sci Med*. 2018;215:123–132.
19. Gelman A, Hill J. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. 1st ed. Cambridge, UK: Cambridge University Press; 2006.
20. Park T, Casella G. The Bayesian lasso. *J Am Stat Assoc*. 2008;103(482):681–686.
21. Carvalho CM, Polson NG, Scott JG. The horseshoe estimator for sparse signals. *Biometrika*. 2010;97(2):465–480.
22. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat Comput*. 27(5):1413–1432.
23. R Foundation for Statistical Computing. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2020.
24. Su YS, Yajima M. R2jags: Using R to Run'JAGS'. 2020. R package version 0.6-1. https://cran.r-project.org/web/packages/R2jags/index.html. Accessed September 22, 2022.

25. Plummer M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, 2003;124(125.10):1–10.
26. Substance Abuse and Mental Health Services Administration. *Behavioral Health Barometer: Ohio, Volume 5: Indicators as Measured Through the 2017 National Survey on Drug Use and Health and The National Survey of Substance Abuse Treatment Services*. Rockville, MD: Substance Abuse and Mental Health Services Administration, (HHS Publication No. SMA-19-Baro-17-OH); 2019.
27. Barocas JA, White LF, Wang J, et al. Estimated prevalence of opioid use disorder in Massachusetts, 2011-2015: a capture-recapture analysis. *Am J Public Health.* 2018; 108(12):1675–1681.
28. Substance Abuse and Mental Health Services Administration. *Behavioral Health Barometer: United States, Volume 5: Indicators as Measured Through the 2017 National Survey on Drug Use and Health and The National Survey of Substance Abuse Treatment Services*. Rockville, MD: Substance Abuse and Mental Health Services Administration, (HHS Publication No. SMA-19-Baro-17-US); 2019.
29. Larochelle MR, Bernstein R, Bernson D, et al. Touchpoints - opportunities to predict and prevent opioid overdose: a cohort study. *Drug Alcohol Depend*. 2019;204:107537.
30. Barocas JA, Wang J, Marshall BDL, et al. Sociodemographic factors and social determinants associated with toxicology confirmed polysubstance opioid-related deaths. *Drug Alcohol Depend*. 2019;200:59–63.
31. Walley AY, Bernson D, Larochelle MR, et al. The contribution of prescribed and illicit opioids to fatal overdoses in Massachusetts, 2013–2015. *Public Health Rep.* 2019;134(6):667–674.
32. Wakeman SE, Larochelle MR, Ameli O, et al. Comparative effectiveness of different treatment pathways for opioid use disorder. *JAMA Netw Open*. 2020;3(2):e1920622.